Morphology renders homophonous segments phonetically different: Word-final /s/ in German

Dominic Schmitz[1], Dinah Baer-Henney[2]

[1]Department of English and American Studies, English Language and Linguistics, Heinrich Heine University Düsseldorf
[2]Department of Linguistics, Heinrich Heine University Düsseldorf

dominic.schmitz@uni-duesseldorf.de, dinah.baer-henney@uni-duesseldorf.de

Abstract

Recent research challenges established models of speech production by revealing unexpected phonetic differences in phonologically identical elements induced by morphological structure. While established models assume that morphology does not play a role in later production stages [1, 2], it has been shown that English word-final /s/ duration is longest in non-morphemic contexts, shorter with suffixes, and shortest in clitics [3, 4]. Subsequent research found that such differences are not only produced but also perceived by listeners and able to influence comprehension [5]. Recently, [6] investigated if German speakers could use subphonemic durational cues to acquire the morphological categories singular (short word-final /f/) and plural (long word-final /f/) in an artificial language. However, the study revealed subphonemic cues were insufficient. Building on English findings, the present production study examines subphonemic durational differences in German word-final /s/. Preliminary results (20 speakers, 823 data points) show significant differences between non-morphemic and plural /s/ duration ($p < 0.0001$; Cohen's $d = 0.3$). The findings challenge established models, suggest that [6]'s null-results were due to the reversed direction of durational cues used, and indicate morphological influences on speech production extend beyond English.

Index Terms: speech production, prosody morphology interface, homophony, subphonemic detail

1. Introduction

A growing number of studies on the acoustic properties of homophonous elements has demonstrated unexpected effects of higher level categories on their phonetic realization. Such effects were found for lexemes [7], for stems [8], for prefixes [9], and, most prominently, for word-final /s/ in English. Here, it was found that non-morphemic /s/ is longest in acoustic duration, followed by suffix /s/, which in turn is followed by clitic /s/ [3, 4]. The present study investigated a similar case in German, motivated by three main arguments.

First, evidence for an influence of higher levels of speech production, e.g., morphology, on the outcome of lower levels of speech production, e.g., prosody, is an important issue for current theories of language production. Most traditional theories of a feed-forward structure [1, 2] do not allow this type of interaction, as homophones should be truly homophonous at the phonological level already. Only few theories allow effects of morphology on phonetics [10, 11], however, only one of these theories – the discriminative lexicon – has implementations available to test such effects [11] and indeed it could be shown that differences in fine phonetic detail can emerge from the mental lexicon [4]. The present study aims at investigating yet another case of possible interplay between the levels of morphology and prosody.

Second, effects of morphology on production are, thus far, predominantly reported for English [3, 12, 4] and, in only a few studies, for Dutch [8, 13]. However, as the consequences for language production models are crucial if such effects truly exist, the field is still in need of more evidence, preferably from languages other than English and Dutch. German is particularly suitable here, as it comes with a pattern comparable to that of English word-final /s/ within its nominal paradigm. Contrary to English, the German plural system is highly variable, and the language displays many ways to pluralize nouns. Still, the -s suffix is one of them. In German, word-final /s/ can function as plural suffix just as it can be part of the stem. There are singular forms ending in /s/ such as *Fuchs* 'fox' or *Gips* 'plaster', and there are plural forms ending in /s/ such as *Jobs* 'job-pl' or *Loks* 'locomotive-pl'. Hence, we are dealing with a non-morphemic and a plural word-final /s/, comparable to English, making German the ideal test case for the investigation of potential durational differences due to morphological categories.

Third, a recent study [6] investigated whether an artificial pattern similar to English word-final /s/ is learnable by German adults. In this artificial language learning study, two groups of subjects failed to learn the differentiation of morphological categories by means of durational differences, while another group of subjects succeeded when categories were distinguished by phonemic differences. In the first two groups, the artificial language decoded grammatical number by durational differences in word-final /f/, and in the third group, the artificial language decoded grammatical number by phonemic differences (word-final /f/ and /p/). Crucially, the phonetic pattern applied required the singular to have the shorter /f/ and the plural to have the longer /f/. If the present study showed that German word-final /s/ in singular and plural forms patterns similar to English word-final /s/, this could be an explanation for the subject's poor performance: It would mean that [6]'s subjects were asked to learn a pattern counter to their L1 experience, and counter to how English looks like, where /s/ in singular forms is longer than /s/ in plural forms and not the other way round. This confound could have inhibited successful learning in German adults.

2. Methods

We performed a pseudoword production study with German adult speakers. The core idea is to show whether fine-phonetic durational differences between different types of /s/ exist in German. If such differences are robust, then they should also be transferred to pseudowords, as was shown by [4]. In addition, pseudowords rule out confounding effects that real words are

known to introduce to experiments: storage effects [14], frequency effects [7], and effects of lexical relatedness [15]. In particular, we compared durations of word-final /s/ in singular forms (non-morphemic /s/) and plural forms (plural /s/). Target pseudowords with word-final /s/ were embedded in real sentence contexts. Across participants, the same target pseudoword occurred as singular or as plural form, see Figure 1.The details of the production experiment are described in the following subsections.
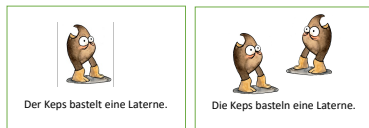

Figure 1: *Comparison between singular forms (non-morphemic /s/, engl. 'The keps makes a lantern.') and plural forms (plural /s/, engl. 'The keps make a lantern.')*

2.1. Speech Materials

As target pseudowords, we designed items whose phonological makeup adhered to possible singular and plural forms in German. Each target ended in a word-final /s/. We used target items with either one or two syllables.

Monosyllabic target items had a /C(C)VCs/ structure. These types occur both as singular and plural forms in German, e.g., *Fuchs* 'fox' vs. *Jobs* 'job-pl' [16]. In these forms, the nucleus is a lax vowel, so we chose vowels from the set {ɑ, ɛ, ɪ, ɔ, ʊ, ʏ, œ}. Before the word-final /s/ we distributed {p, t, k} as first coda consonant evenly across the item set. The onset consonant was chosen such that the target was a pseudoword. Disyllabic target items had a /C(C)V.CVCs/ structure. Again, these types occur both as singular and plural forms in German, e.g., *Rotfuchs* 'red fox-pl' vs. *Bisons* 'bison-pl'. The first syllable of the two syllable target consisted of a (complex) onset and a tense vowel from the set {a, e, i, o, u, y, ø} as nucleus to attract stress. The structure of the second syllable was based on the restrictions that also applied in the monosyllabic targets. Overall, we created a list with $n = 21$ monosyllabic and $n = 21$ disyllabic target items (each of which comes with an even distribution of the 7 vowels x 3 coda consonants).

As additional filler items, we included 10 disyllabic pseudowords ending in -/ən/ to have relatively unequivocal plural forms [17] and 11 disyllabic pseudowords with various shapes used as singular forms. All experimental items conformed to German phonotactics [18].

Target and filler items were embedded in real German sentences. The sentences were created such that they describe daily activities. The sentences followed the simple structure subject – verb – object. The target pseudoword appeared as an alien creature - the noun in the sentence's subject. The subject consisted of a determiner and a noun, forming a nominal phrase. For determiners of singular forms, we alternated between *der* 'the-m.sg' and *das* 'the-n.sg'. For plural forms, we used *die* 'the-pl'. The target pseudoword with word-final /s/ was followed by the predicate, a verb form in present tense. We used transitive verbs only and added a plausible real animate or inanimate object, e.g., *Der/Das/Die X liebt/lieben Süßigkeiten.* 'The X love(s) sweets.'. To facilitate isolating the target sound in the analyses, we chose that the following verb form never began with a fricative. The content of each sentence was plausible

for both subjects in the singular and in the plural. Grammatical number was unambiguous for all nouns as the determiner or, if not the determiner, the subject-verb-agreement in the sentence was disambiguating.

2.2. Procedure

Subjects were instructed that they would get to know alien creatures (we used alien pictures by [19]) and that these aliens had unfamiliar names but engage in daily-life activities just as humans. As an explanation for the simplicity of the setup, we argued that adult subjects were part of a control group for a study with children. The trials followed the following procedure, see Figure 2. Each alien/item was presented in three consecutive slides. Subjects were able to go through the presentation in a self-paced manner. First, subjects saw a picture and the name of the alien creature, then a brief explanation of the situation. Finally, a question was asked. The subjects' task was to answer the question aloud with the situation sentence they saw before. We recorded their answers. The experiment was implemented in WikiSpeech [20] and run online.

Figure 2: *Procedure: three example slides of one trial that were presented in consecutive order, English translation: 'This is a keps.', 'The keps makes a lantern.', 'What does the keps do?'*

2.3. Hypotheses

Following established theories of speech production [1, 2] and the findings of previous research on word-final /s/ in English, the present study investigated the following two hypotheses:

- Hypothesis 1: Feed-forward hypothesis
  There are no durational differences between non-morphemic and plural /s/.

- Hypothesis 2: Morphological effect hypothesis
  There are durational differences between non-morphemic and plural /s/.

2.4. Subjects

Twenty native German adults took part in the experiment and, following the study by [4], we plan to test and analyze at least another 20 speakers. The current speakers have a mean age of 29.6 years and cover an age range of 20 to 61 years.

2.5. Analyses

The acoustic analyses of all recordings took place in Praat [21]. For all items, the duration of /s/ was determined using both the spectrogram and the waveform, following segmentation criteria based on the phonetic literature [22]. All boundaries were placed at the nearest zero crossing. Recordings with production errors, stutter, and laughter were excluded. Overall, 823 of 840 recordings were retained for further analysis. Based on the annotations in Praat, the durations of the recorded /s/ sounds were extracted using the rPraat package [23] in R [24].

Further variables specified for each trial were TYPEOFS (non-morphemic or plural), number of syllables (NSYLL),

phonological neighborhood density ([25]; NEIGHDEN), speaking rate (using a Praat script by [26]; RATE), the segment preceding the word-final /s/ (PREC), the segment following the word-final /s/ (FOLC), pause duration if the item was followed by a pause (PAUSEDUR), and the pertinent ITEM. To account for inter-speaker differences, speaker AGE, additional languages (ADDITIONALLS), and the SPEAKER ID were included.

The duration of /s/, SDUR, and speaking rate were log-transformed due to their non-normal distributions. The log-transformed variables are SDURLOG and RATELOG. SDUR-LOG follows a normal distribution (Shapiro-Wilk test: $p = 0.1079$), while RATELOG is still non-normally distributed, but less so than RATE (Shapiro-Wilk tests: $p = 5.1e^{-09}$ vs. $p = 2.2e^{-16}$).

To avoid issues of collinearity, which might lead to unreliable model estimates [27], all predictor variables were checked for high levels of correlation (i.e., $|\rho| \geq 0.4$). Strong correlations were found for TYPEOFS and PREC ($\rho = -0.42$) and for NSYLL and NEIGHDEN ($\rho = -0.44$). We excluded PREC to retain the predictor of interest, i.e. TYPEOFS, and excluded NSYLL, as NEIGHDEN is a more fine-grained measure.

Finally, SDURLOG entered a linear mixed-effects regression analysis as dependent variable using the packages lme4 [28] and lmerTest [29] in R. TYPEOFS, PREC, FOLC, NEIGHDEN, RATELOG, AGE, and PAUSEDUR entered the model as fixed effects. ITEM, ADDITIONALLS, and SPEAKER were specified as random intercepts. After step-wise model reduction by removing variables that do not improve model fit, the final model contained TYPEOFS, FOLC, and PAUSEDUR as fixed effects and SPEAKER as random intercept.

3. Results

The predictor of interest, TYPEOFS, showed a significant effect on log-transformed /s/ durations. Its effect is shown in Figure 3. Non-morphemic /s/ comes with significantly longer durations than plural /s/. The observed effect size, as measured by Cohen's d, was $d = 0.3$ [30]. The results of all fixed effects are given in Table 1.

Table 1: *p-values of fixed effects in the final model, fitted to the log-transformed durations of /s/.*

Fixed effect	DenDF	F-value	*p*-Value
TYPEOFS	803.06	16.21	< 0.0001
FOLC	803.11	8.68	< 0.0001
PAUSEDUR	821.07	28.91	< 0.0001

4. Discussion

The present study set out to investigate whether there are durational differences between word-final non-morphemic and plural /s/ in German. Speech data elicited in an online production experiment was analyzed following previous research on word-final /s/ in English. It was found that in German, just like in English, non-morphemic /s/ shows longer acoustic durations than plural /s/. Hence, hypothesis 2, the morphological effect hypothesis, is confirmed: There are durational differences between non-morphemic and plural /s/. This finding comes with important consequences for the three main arguments motivating the present research.
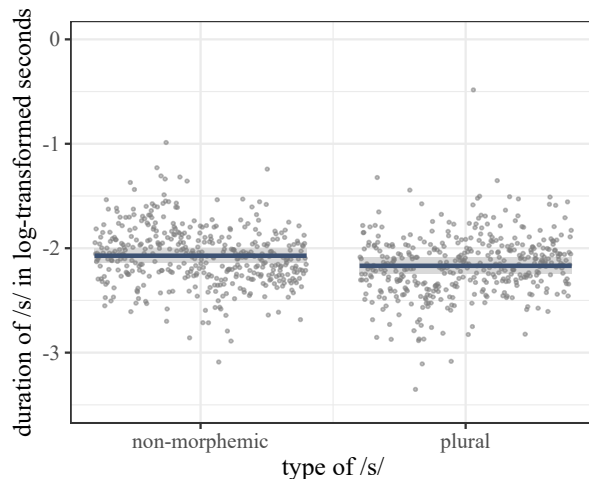

Figure 3: *Effect of type of /s/ (non-morphemic vs. plural) in the linear mixed-effects regression model, fitted to the log-transformed values of duration of /s/*

Interestingly, the durational differences between non-morphemic and plural /s/ patterns similarly in both English and German, with non-morphemic /s/ being longer than plural /s/. [31] showed that the differences between English non-morphemic and plural /s/ are, in part, explained by their resonance within the mental lexicon. They found that, among other factors, the degree of semantic co-activation, the semantic similarity to other entries, and the certainty in a word's form are able to explain /s/ durations. If future research shows that similar factors account for the differences found in the present study for German word-final /s/, established theories of speech production do not only need to include morphological effects into their account of phonetics, but should also consider semantic and psycholinguistic factors. An alternative explanation may be a different representation of non-morphemic and plural forms under the account of Prosodic Phonology [32, 33]. As a consequence, durational differences would result from a difference in the prosodic structure of these forms. While it may be debatable whether a plural /s/ needs to be represented as an affixal or an internal clitic within the prosodic hierarchy [34, 35], it is not controversial that a non-morphemic /s/ is located more central in the prosodic hierarchy [36] than any type of representation of the plural /s/. However, the theory does not make concrete predictions about the direction of the effect. One valid assumption would indeed be that more peripheral units may be expected to be longer than those located more centrally, while we observe the opposite.

Adding German to the rather short list of languages for which morphologically induced phonetic differences were attested increases the pressure on established theories of speech production, which do not assume any influence of morphology on prosody. The present results demonstrate that morphology does indeed influence phonetic realizations, rendering a strict separation of morphology and phonetics extremely unlikely. While we provide evidence from German, we acknowledge that core assumptions on the architecture of language production models should be backed up with evidence from other, more diverse and preferably non-Germanic languages.

For the learning study by [6], our results indicate what was foreshadowed earlier: Subjects were asked to learn a pattern

counter to their language experience. While in German word-final /s/ we found that non-morphemic /s/ showed longer durations than plural /s/, the authors of the learning study asked L1 speakers of German to learn the distinction between singular and plural in an artificial language in which the plural suffix /f/ showed a longer acoustic duration than its non-morphemic counterpart. With the subphonemic durational pattern in the artificial language contradicting the one in their L1, speakers may not have been able to acquire the number system of the artificial language, the only learning cue being the non-intuitive durational pattern.

5. Conclusions

This study was the first to investigate subphonemic durational differences in word-final /s/ in German. In line with previous results on English word-final /s/, we found that non-morphemic /s/ is longer than plural /s/. Using carefully controlled pseudowords and contexts, we demonstrated that durational differences between different types of /s/ are of a robust nature. This suggests that similar results for English were not a by-product of study design or a result of language-specific features. Instead, we conclude that differences in /s/ durations are due to the morphological categories of the different types of /s/, and, in turn, the semantic and psycholinguistic properties of the /s/-bearing words. Our results call for a revision of established models of speech production.

6. Acknowledgments

7. References

[1] P. Kiparsky, "Lexical morphology and phonology," in *Linguistics in the morning calm: Selected papers from SICOL1*, I.-S. Yang, Ed., 1982, pp. 3–91.

[2] A. Roelofs and V. S. Ferreira, "The architecture of speaking," in *Human language: From genes and brains to behavior*, P. Hagoort, Ed. Cambridge: MIT Press, 2019, pp. 35–50.

[3] I. Plag, J. Homann, and G. Kunter, "Homophony and morphology: The acoustics of word-final S in English," *Journal of Linguistics*, vol. 53, no. 1, pp. 181–216, 2017.

[4] D. Schmitz, D. Baer-Henney, and I. Plag, "The duration of word-final /s/ differs across morphological categories in English: Evidence from pseudowords," *Phonetica*, vol. 78, no. 5-6, pp. 571–616, 2021.

[5] D. Schmitz, *Production, perception, and comprehension of subphonemic detail: Word-final /s/ in English*. Language Science Press, 2022.

[6] D. Baer-Henney and D. Schmitz, "In morphological learning, phonemic cues are stronger than phonetic cues," *Manuscript*, subm.

[7] S. Gahl, "Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech," *Language*, vol. 84, pp. 474–496, 2008.

[8] R. J. Kemps, L. H. Wurm, M. Ernestus, R. Schreuder, and H. Baayen, "Prosodic cues for morphological complexity in Dutch and English," *Language and Cognitive Processes*, vol. 20, no. 1-2, pp. 43–73, 2005.

[9] S. B. Hedia and I. Plag, "Gemination and degemination in English prefixation: Phonetic evidence for morphological organization," *Journal of Phonetics*, vol. 62, pp. 34–49, 2017.

[10] J. B. Pierrehumbert, "Exemplar dynamics," in *Frequency effects and the emergence of linguistics structure*, J. Bybee and P. Hopper, Eds. Amsterdam: John Benjamins, 2001, pp. 137–158.

[11] R. H. Baayen, Y.-Y. Chuang, E. Shafaei-Bajestan, and J. P. Blevins, "The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning," *Complexity*, vol. 2019, p. 4895891, 1 2019.

[12] J. Zimmermann, "Morphological status and acoustic realization," in *Proceedings of the Sixteenth Australasian International Conference on Speech Science and Technology (SST-2016)*. ASSTA, 2016, pp. 201–204.

[13] R. J. Kemps, M. Ernestus, R. Schreuder, and R. Harald Baayen, "Prosodic cues for morphological complexity: The case of Dutch plural nouns," *Memory & cognition*, vol. 33, no. 3, pp. 430–446, 2005.

[14] N. K. Caselli, M. K. Caselli, and A. M. Cohen-Goldberg, "Inflected words in production: Evidence for a morphologically rich lexicon," *Quarterly Journal of Experimental Psychology*, vol. 69, pp. 432–454, 2016.

[15] H. Schriefers, A. D. Friederici, and U. Rose, "Context effects in visual word recognition: Lexical relatedness and syntactic context," *Memory and Cognition*, vol. 26, pp. 1292–1303, 1998.

[16] H. Wegener, "Der Erwerb eines komplexen morphologischen Systems in DaZ: Der Plural deutscher Substantive," in *Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitspracherwerbsforschung. (Linguistische Arbeiten, 520)*. Tübingen: Niemeyer, 2008, pp. 93–117.

[17] K.-M. Köpcke, *Schemata bei der Pluralbildung im Deutschen: Versuch einer kognitiven Morphologie*. Tübingen: G. Narr, 1993.

[18] R. Wiese, *The phonology of German*. Oxford: Oxford University Press, 2000.

[19] R. van de Vijver and D. Baer-Henney, "Developing biases," *Frontiers in Psychology*, vol. 5, p. 634, 2014.

[20] C. Draxler and K. Jänsch, "Wikispeech – a content management system for speech databases," in *Proceedings of Interspeech*, 2008, pp. 1646–1649.

[21] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer," 2019. [Online]. Available: http://www.praat.org/

[22] P. Ladefoged, *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Wiley, 2003.

[23] T. Boril and R. Skarnitzl, "Tools rpraat and mpraat," in *Text, Speech, and Dialogue: 19th International Conference, TSD 2016, Brno, Czech Republic, September 12-16, 2016, Proceedings*, P. Sojka, A. Horák, I. Kopecek, and K. Pala, Eds. Cham: Springer International Publishing, 2016, pp. 367–374.

[24] R. C. Team, "R: A language and environment for statistical computing," 2021. [Online]. Available: https://www.r-project.org/

[25] V. Marian, J. Bartolotti, S. Chabal, and A. Shook, "Clearpond: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities," *PLoS ONE*, vol. 7, p. e43230, 2012.

[26] N. de Jong and T. Wempe, "Praat script syllable nuclei [Praat script]," 2008. [Online]. Available: https://sites.google.com/site/speechrate/Home/praat-script-syllable-nuclei-v2

[27] F. Tomaschek, P. Hendrix, and R. H. Baayen, "Strategies for addressing collinearity in multivariate linguistic data," *Journal of Phonetics*, vol. 71, pp. 249–267, 2018.

[28] D. Bates, M. Mächler, B. M. Bolker, and S. C. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, 2015.

[29] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmertest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, 2017.

[30] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. New York: Routledge, 1988.

[31] D. Schmitz, I. Plag, D. Baer-Henney, and S. D. Stein, "Durational differences of word-final /s/ emerge from the lexicon: Modelling morpho-phonetic effects in pseudowords with linear discriminative learning," *Frontiers in Psychology*, vol. 12, 2021.

[32] G. E. Booij, "Principles and parameters in prosodic phonology," 1983.

[33] M. Nespor and I. Vogel, *Prosodic phonology: with a new foreword*. Walter de Gruyter, 2007, vol. 28.

[34] H. Goad, "Plurals in sli: Prosodic deficit or morphological deficit?" *Language Acquisition*, vol. 7, no. 2-4, pp. 247–284, 1998.

[35] ——, "Markedness in right-edge syllabification: Parallels across populations," *Canadian Journal of Linguistics/Revue canadienne de linguistique*, vol. 47, no. 3-4, pp. 151–186, 2002.

[36] E. Selkirk, "The prosodic structure of function words," in *Signal to syntax*. Psychology Press, 2014, pp. 199–226.